LET IT SNOW, LET IT FLOW: ASSESSING FUTURE CHANGES IN GREENLAND ICE SHEET RUNOFF USING MACHINE LEARNING AND CLIMATE MODELS

ELKE SCHLAGER, PHD STUDENT
AARHUS UNIVERSITY
DANISH METEOROLOGICAL INSTITUTE

DMI

WHAT IS ICE SHEET RUNOFF?

SURFACE MASS BALANCE & RUNOFF

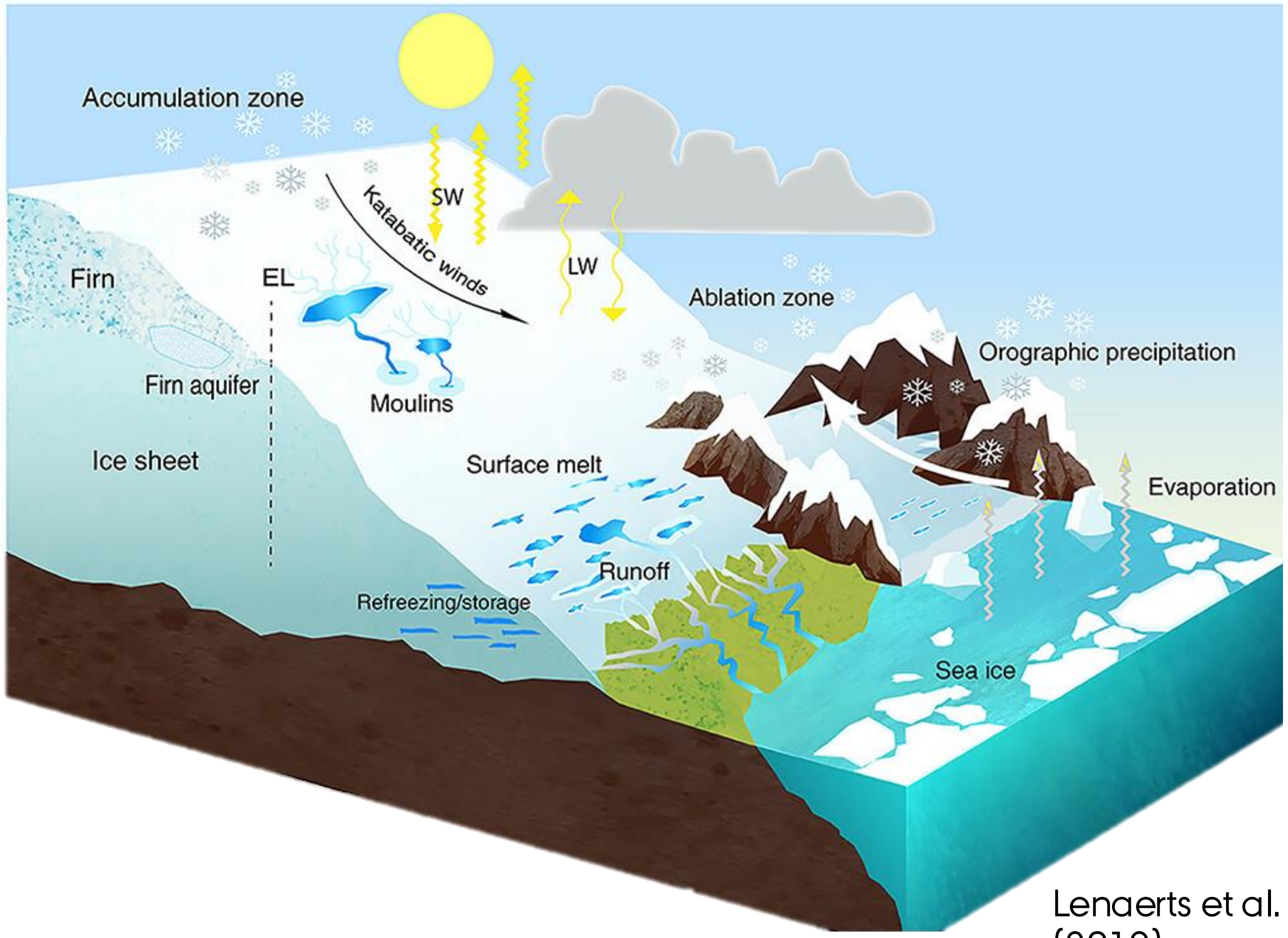
SMB = Precipitation
- Sublimation/Evaporation
- Runoff

Runoff = Rainfall
+ Surface Melt
- Refreezing
- Retention

from atmosphere

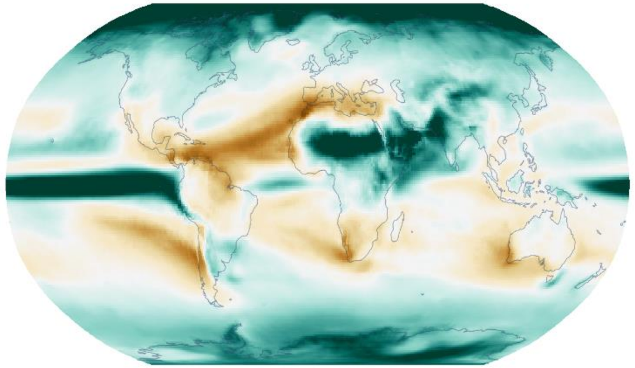
both

within the snow/firn pack

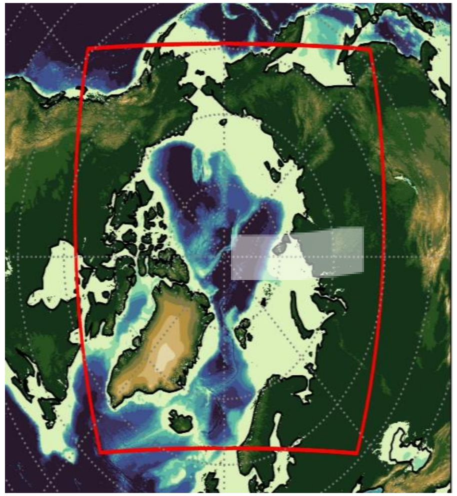


Lenaerts et al. (2019)

POLAR REGIONAL CLIMATE MODELS

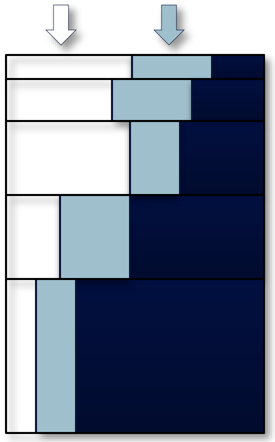


Input: atmospheric data from GCM

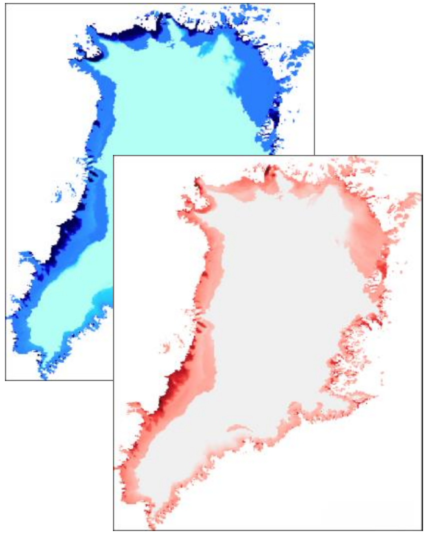


(a) atmospheric model

+



(b) firn model



SMB outputs (surface melt, albedo, runoff, refreezing, …)

WHY USING MACHINE LEARNING?

WHAT DO WE WANT?

RCMs are computationally expensive…

… want to emulate polar RCMs to

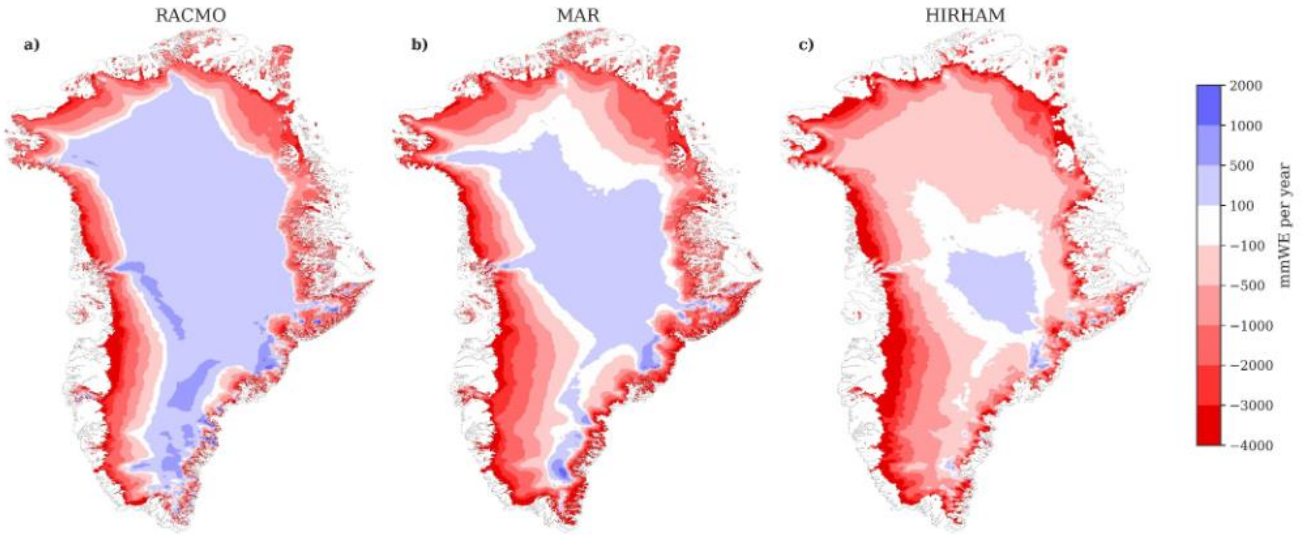
→ predict daily runoff

→ build large ensembles for SMB

→ explore dynamic topography

→ run sensitivity studies

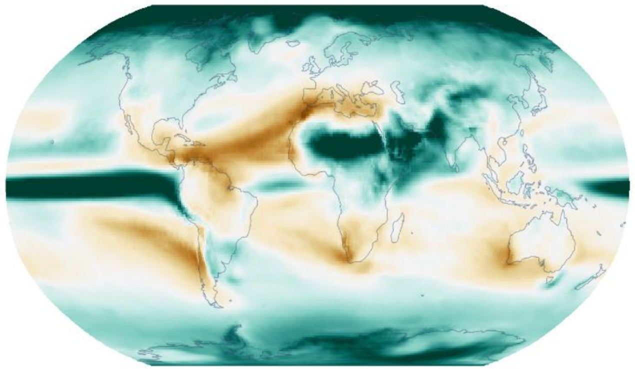
→ integrate into ESMs



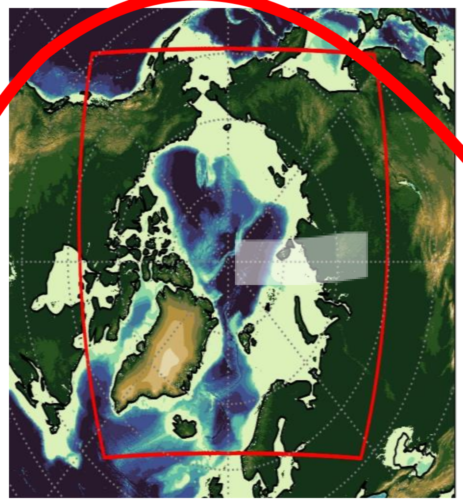
Discrepancies in SMB projections from different polar RCMs (Glaude et al. 2024)

HOW TO USE MACHINE LEARNING?

POLAR REGIONAL CLIMATE MODELS

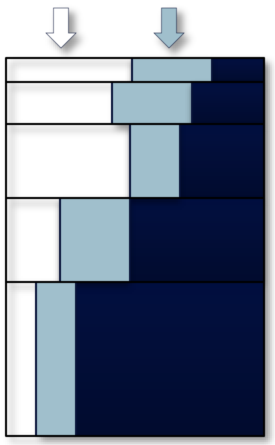


Input: atmospheric data from GCM



(a) atmospheric model

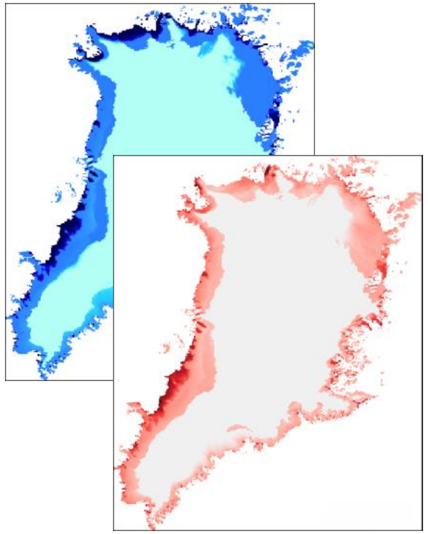
\+



(b) firn model

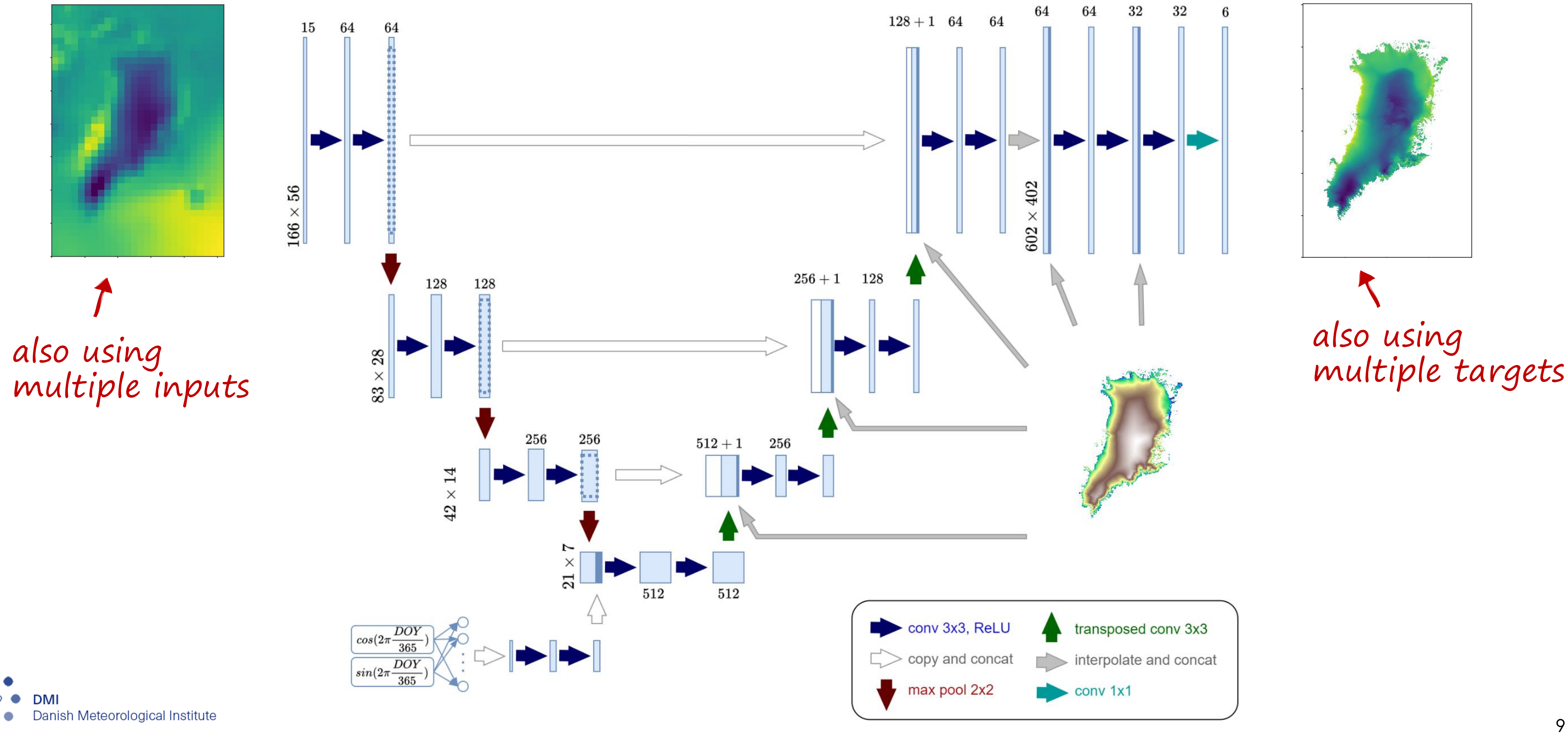
for ML: a spatial problem!

for ML: a temporal problem!



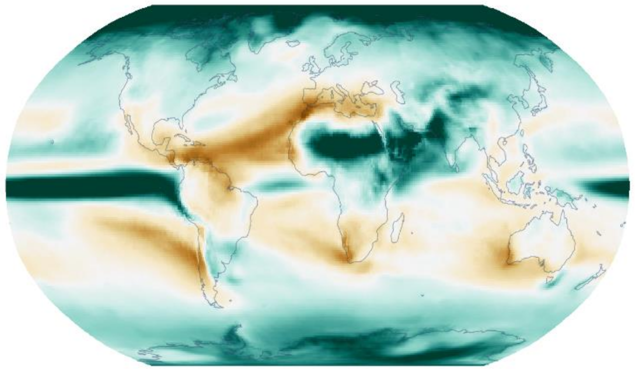
SMB outputs (surface melt, albedo, runoff, refreezing, ...)

UNET FOR ATMOSPHERIC DOWNSCALING

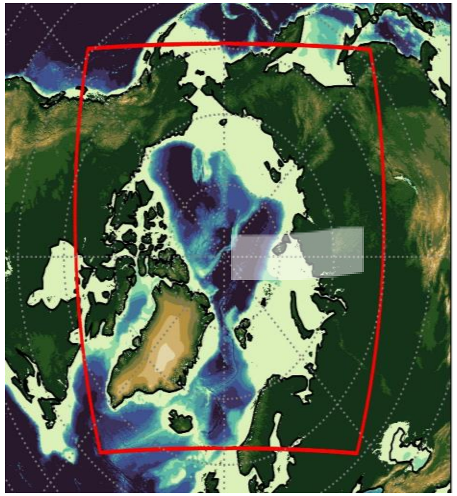


DMI
Danish Meteorological Institute

POLAR REGIONAL CLIMATE MODELS

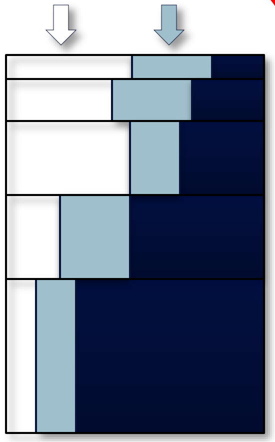


Input: atmospheric data from GCM

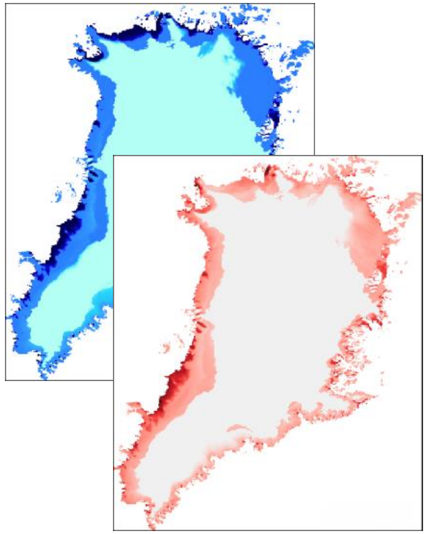


(a) atmospheric model

+



(b) firn model

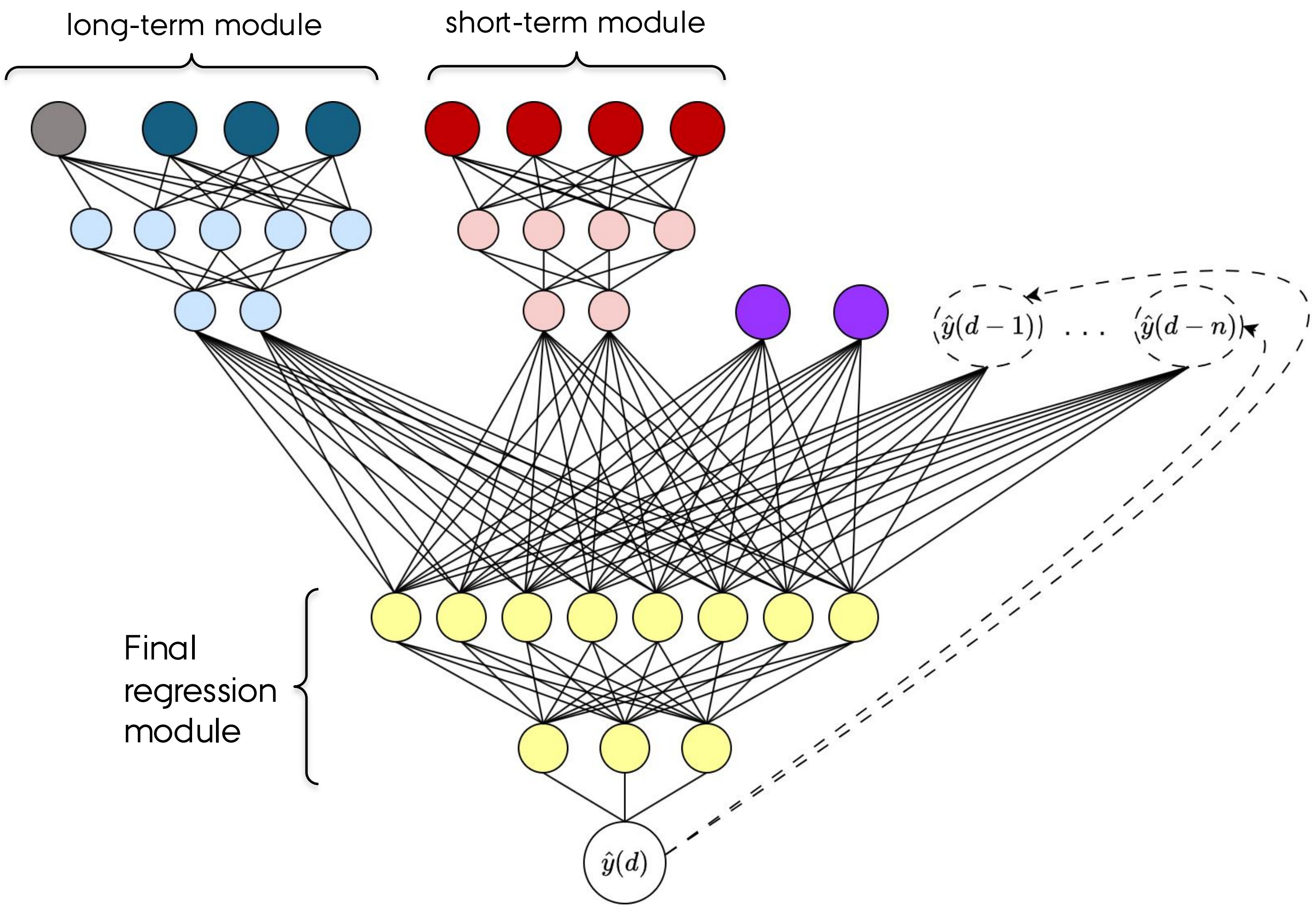


SMB outputs (surface melt, albedo, runoff, refreezing, ...)

for ML: a spatial problem!

for ML: a temporal problem!

A MODULAR NEURAL NETWORK



CHALLENGES

Which variables?

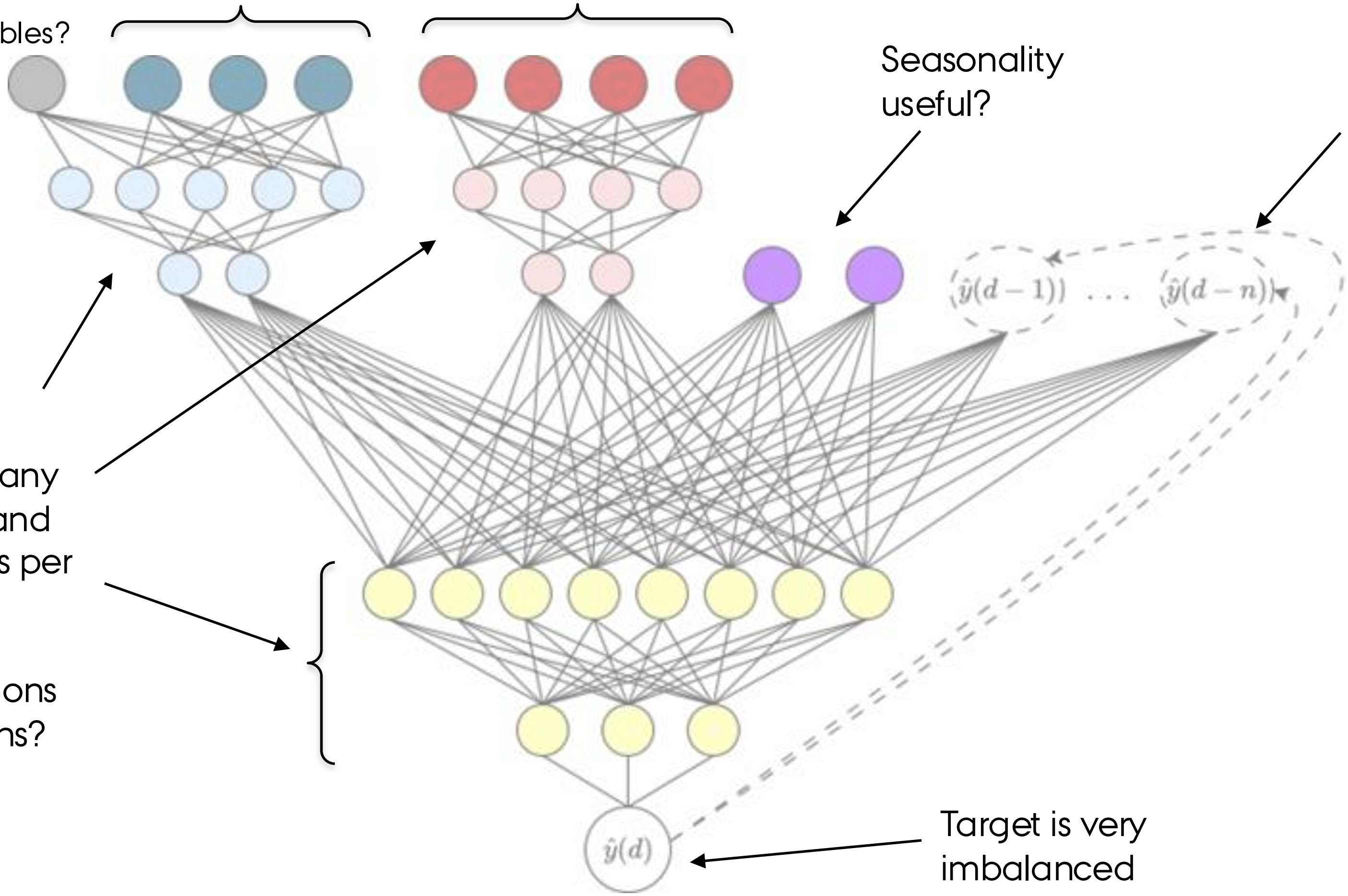
Which variables?
How many months?

Which variables?
How many days?

Some inputs are highly imbalanced and skewed;
some data seems unreasonable

Seasonality useful?

How many auto-regressive steps?



How many layers and neurons per layer?
Which activations functions?

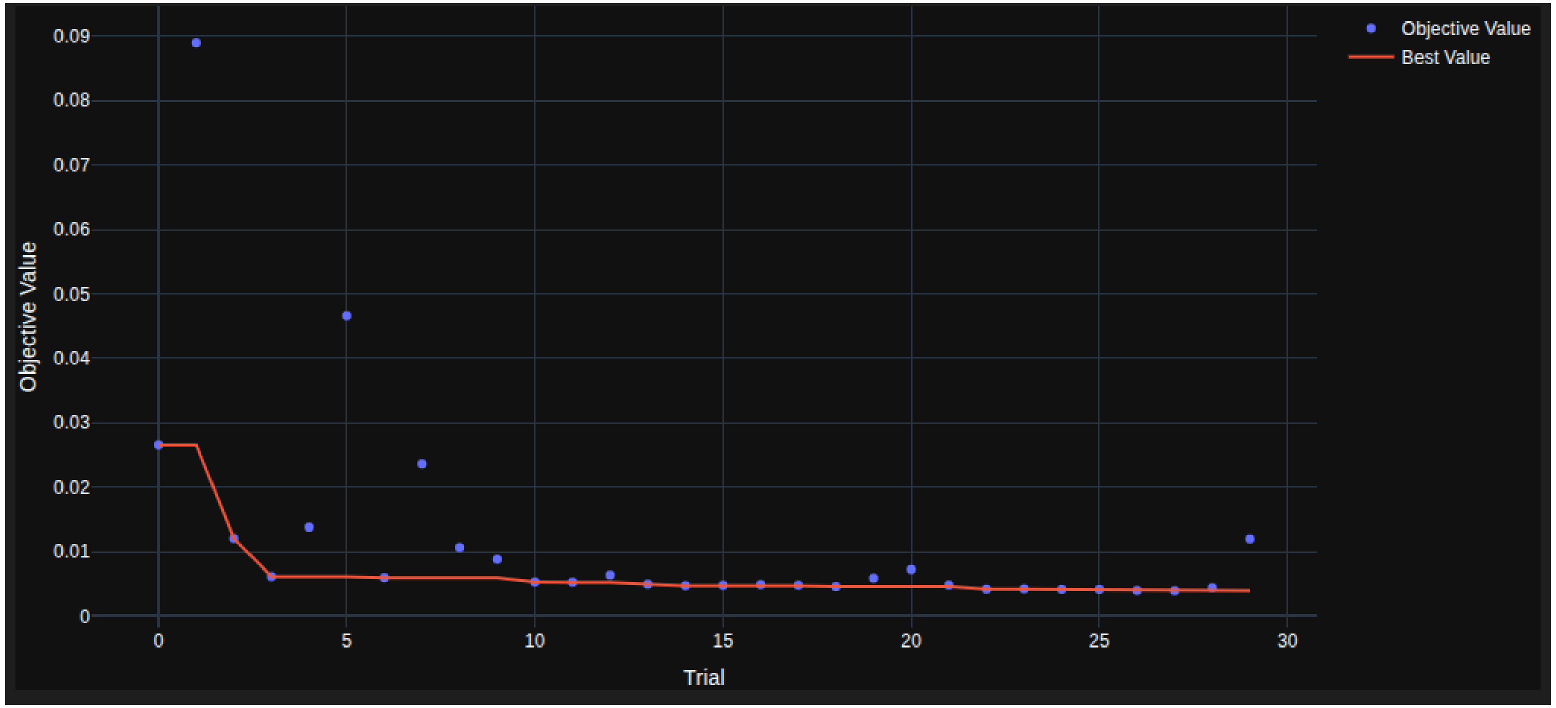
Target is very imbalanced

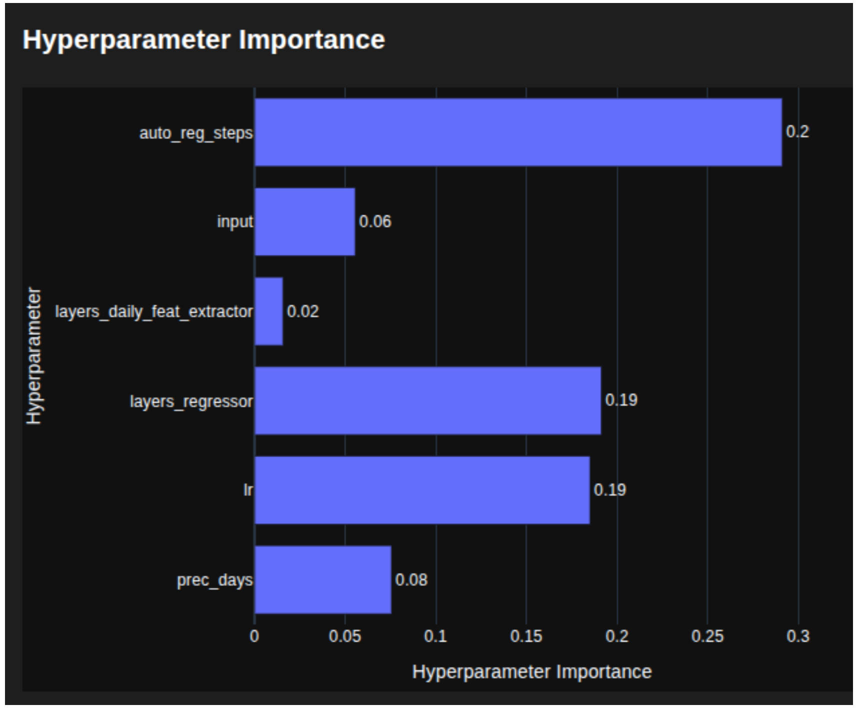
Large amount of data – need to optimize:
→ File storage
→ Data reading pipeline
→ Training loop

Zarr
EUROPEAN WEATHER CLOUD
PyTorch
OPTUNA

HYPERPARAMETER OPTIMIZATION

OPTUNA – A Framework for Hyperparameter tuning using Bayesian Optimization



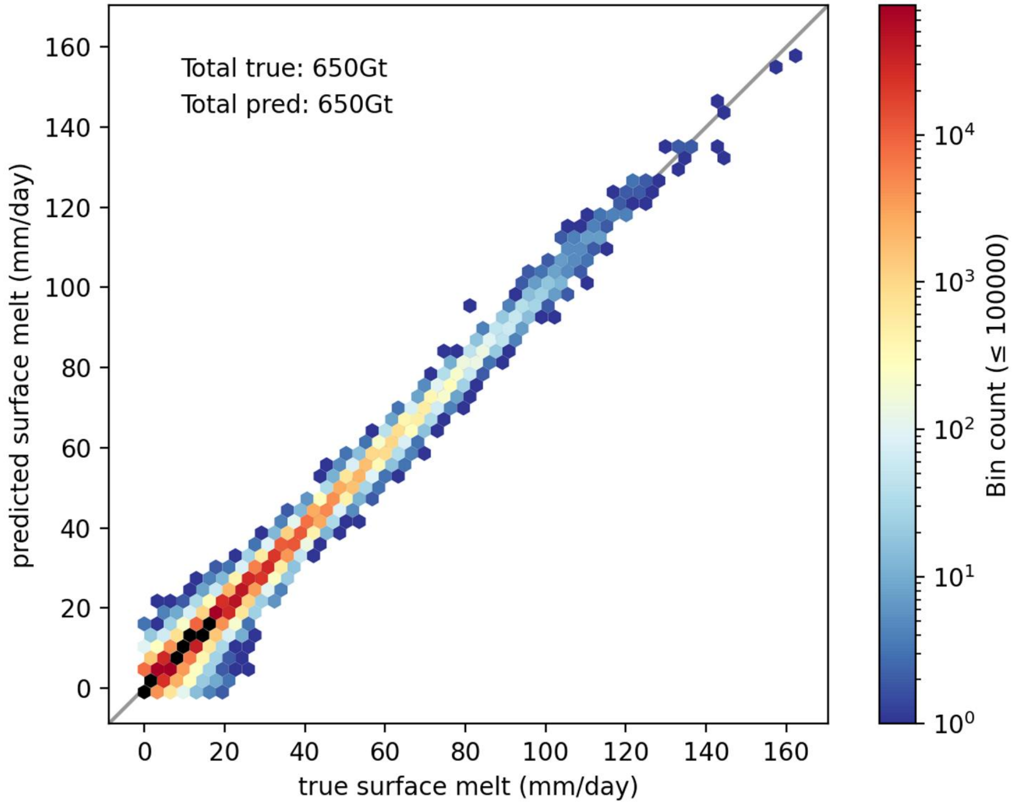


RESULTS SO FAR . . .

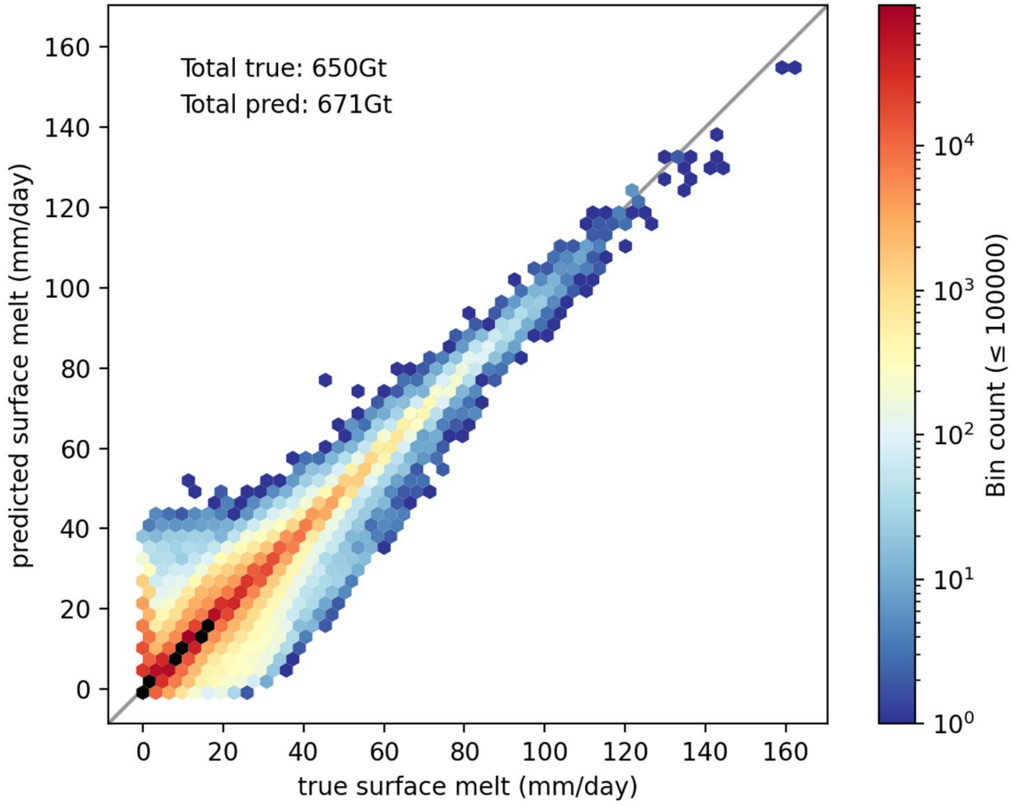
RESULTS FOR SURFACE MELT

Model with 5 days of auto-regressive steps → a lot of dependence on the previous predictions
→ error propagation

When using true previous melt:

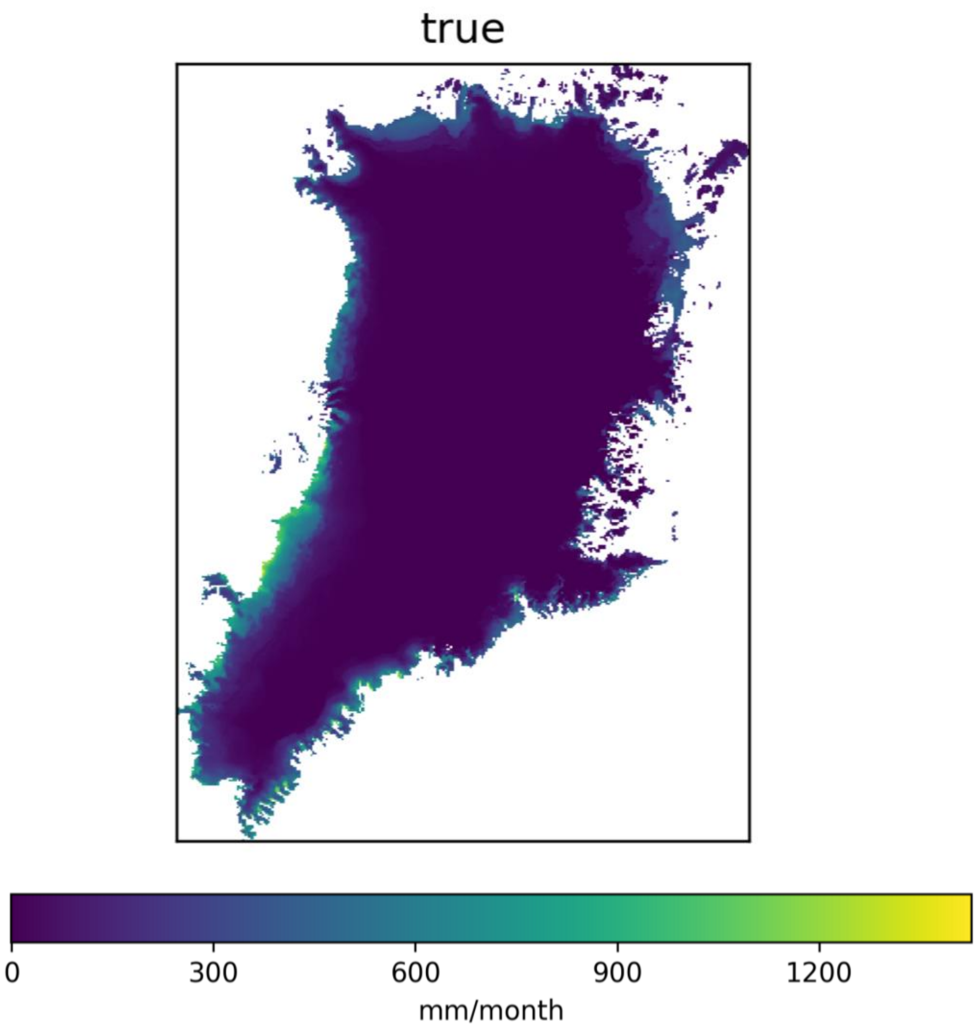


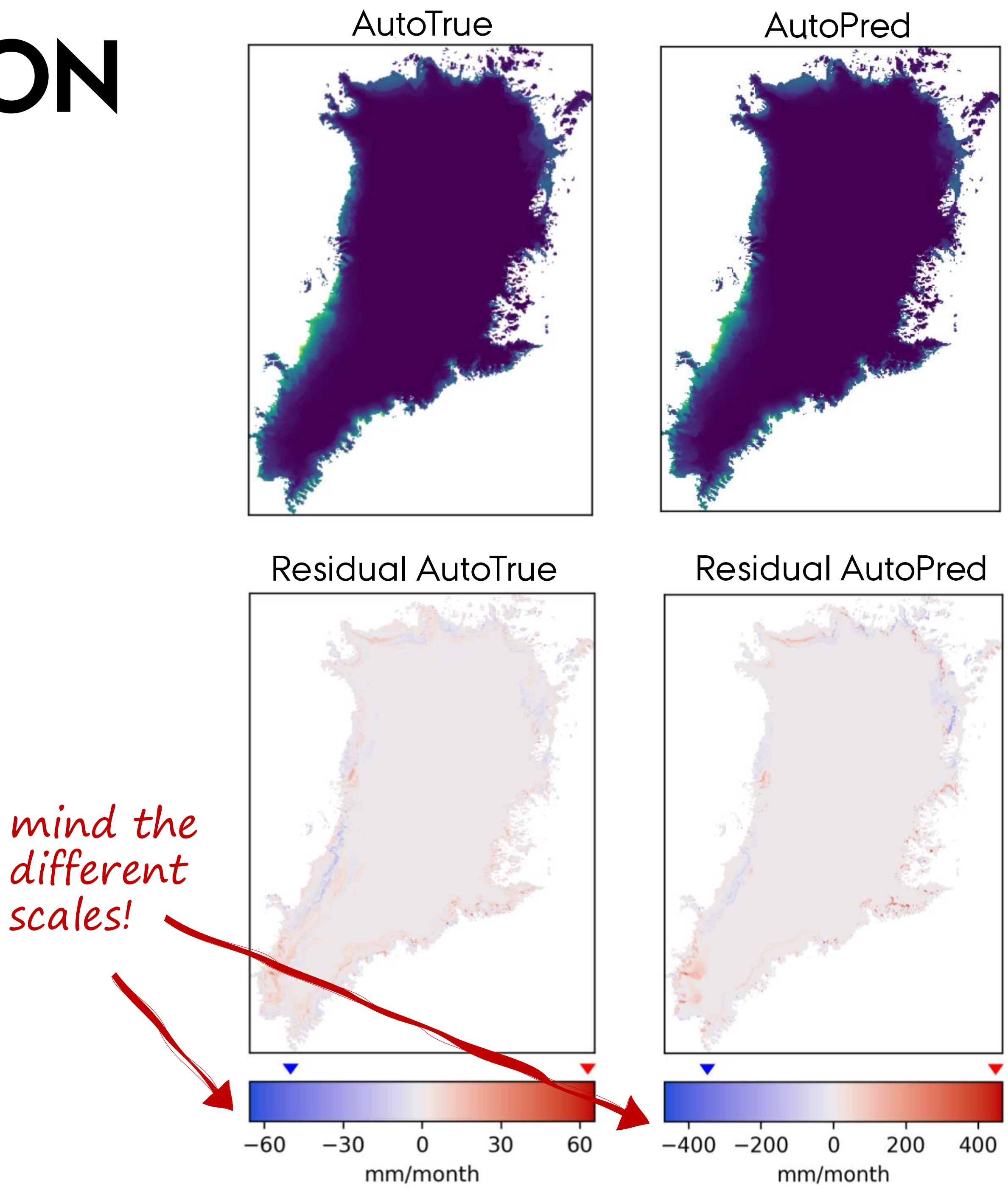
When using previous melt predictions:



ERROR-PROPAGATION

Temporal aggregation for August





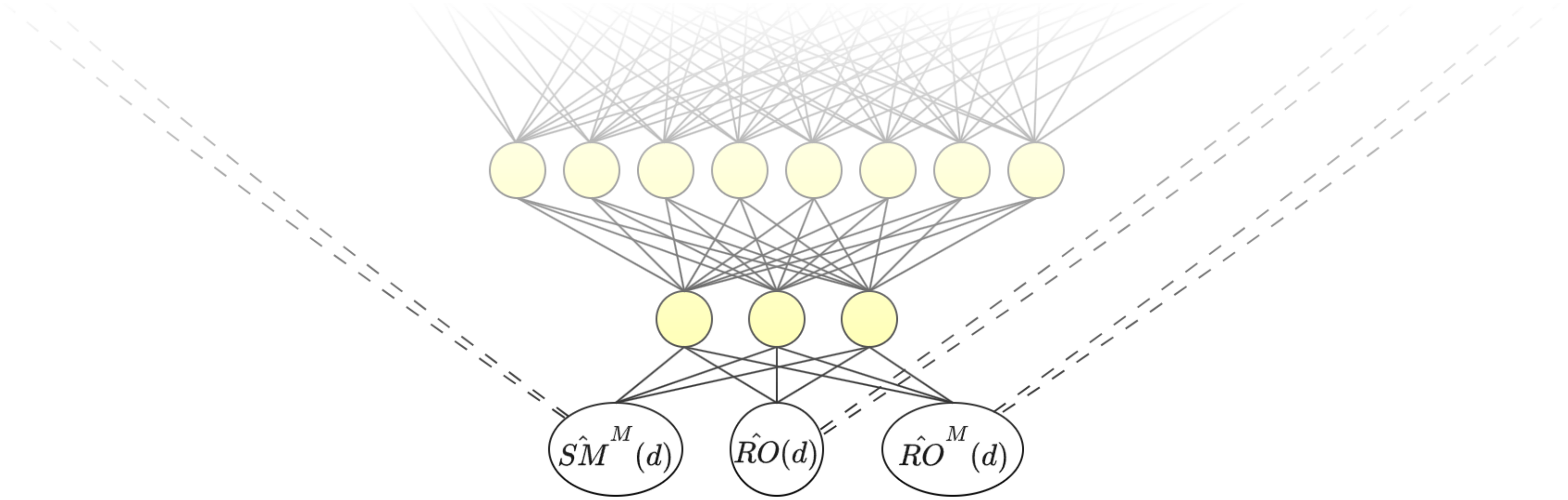
ONGOING WORK

MULTI-TARGET REGRESSION

- Predict surface melt and runoff simultaneously
- hybrid loss for classification and regression

How to choose those?

$$\mathcal{L} = \alpha \cdot BCE\left(\widehat{SM}^M, SM^M\right) + \beta \cdot MSE_{RO^M=1}\left(\widehat{RO}, RO\right) + \gamma \cdot BCE\left(\widehat{RO}^M, RO^M\right)$$



MULTI-TARGET REGRESSION

$$\mathcal{L} = \frac{1}{\sigma_1^2} \cdot BCE\left(\widehat{SM}^M, SM^M\right) + \log(\sigma_1)$$

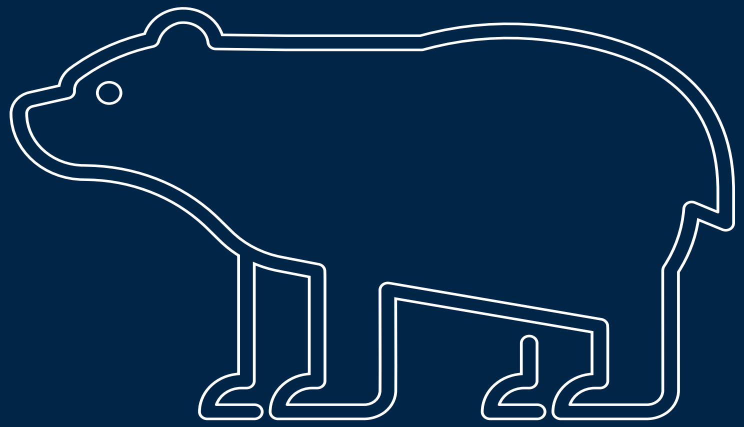
$$+ \frac{1}{2\sigma_2^2} \cdot MSE_{RO^M=1}\left(\widehat{RO}, RO\right) + \log(\sigma_2)$$

$$+ \frac{1}{\sigma_3^2} \cdot BCE\left(\widehat{RO}^M, RO^M\right) + \log(\sigma_3)$$

Using Aleatoric homoscedastic uncertainty (Kendall et al. 2018)

EVALUATION STRATEGIES

- Evaluate spatially integrated error and bias
- Evaluate per zones and basins
- Evaluate extremes
- Evaluate error propagation
- Feature Importance and Explainability ?



ELKE SCHLAGER

MSc, PhD student

eschlager@envs.au.dk

Aarhus University & Danish Meteorological Institute